Acta Biochimica Polonica

Vol. 52 No. 4/2005, 943–946

on-line at: www.actabp.pl

Communication

Isolation and expression pattern of *RGS21* gene, a novel RGS member

Xin Li, Lei Chen, Chaoneng Ji, Bing Liu, Jiefeng Gu, Jian Xu, Xianqiong Zou, Shaohua Gu and Yumin Mao✉

State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Fudan University, Shanghai, People's Republic of China; ✉e-mail: ymmao@fudan.edu.cn

Received: 01 December, 2004; revised: 13 June, 2005; accepted: 30 June, 2005
available on-line: 08 September, 2005

Regulators of G-protein signaling (RGS) proteins are known for the RGS domain that is composed of a conserved stretch of 120 amino acids, which binds directly to activated G-protein α subunits and acts as a GTPase-activating protein (GAP), leading to their deactivation and termination of downstream signals. In this study, a novel human RGS cDNA (*RGS21*), 1795 bp long and encoding a 152-amino acid polypeptide, was isolated by large-scale sequencing analysis of a human fetal brain cDNA library. Unlike other RGS family members, *RGS21* gene has no additional domain/motif and may represent the smallest known member of RGS family. It may belong to the B/R4 subfamily, which suggests that it may serve exclusively as a negative regulator of αi/o family members and/or αq/11. PCR analysis showed that *RGS21* mRNA was expressed ubiquitously in the 16 tissues examined, implying general physiological roles.

Keywords: *RGS21*, G-protein α subunit (Gα), G-protein signaling pathway

Regulators of G-protein signaling (RGS) proteins contain the RGS domain that is composed of a conserved 120 amino acid chain (De Vries *et al*., 1995). The RGS domain acts as a GTPase-activating protein (GAP) that reduces the signal transmitted by the receptor-activated (GTP-bound) G-alpha (Gα) subunit by rapidly returning it to the inactive (GDP-bound) state (Berman *et al*., 1996). Resolution of the crystal structure of the RGS4 protein complexed with a stable transition state mimic of Gα-GTP has revealed that the RGS domain forms nine α-helices that fold into two small subdomains. These subdomains each contact the Gα surface at three distinct sites (Tesmer *et al*., 1997). RGS proteins are found in eukaryotic species ranging from yeast to mammals and are a family of highly diverse, multifunctional signaling proteins. Apart from the RGS domain, RGS proteins differ widely in their overall size and amino-acid identity, and possess a remarkable variety of structural domains and motifs. These additional domains, like DEP, PDZ, PH and so on, link the RGS proteins to other members of the signaling network, where they constitute effector-type molecules (Hollinger & Hepler, 2002). Based on amino-acid sequence similarity, RGS proteins are divided into six subfamilies: RZ, R4, R7, R12, RA and RL. Unlike members of the other four subfamilies possessing at least two domains/motifs, the members of RZ and R4 have only an RGS domain, but the members of RZ have an additional cysteine string compared to R4 members (Zheng *et al*., 1999).

On the basis of sequence similarity, Gα subunits have been divided into four subfamilies: Gαi/o, Gαq/11, Gαs and Gα12/13 (Simon *et al*., 1991). As a general rule, nearly all RGS family members tested act selectively as GAPs for one or more members of the Gαi/o and/or Gαq/11, but not Gαs and Gα12/13, by which the RGS proteins lead to the modulation of adenylate cyclase activity, inhibition of mitogen activated protein kinase (MAPK) and inositol (1,4,5)-trisphosphate/Ca^{2+} signaling, and others (Hollinger & Hepler, 2002). In addition, more and more reports show that RGS proteins associate directly with many diseases and that together with G protein-coupled receptors they are ideal drug targets (Jones *et al*., 2000; Zhong & Neubig 2001). It is apparent that studies on the cloning of RGS genes and exploring their functions are of significance.

During large-scale sequencing we have isolated a novel full-length cDNA from a human fetal brain library, containing an RGS domain. Due to its notable RGS domain, we termed it regulators of G-

Abbreviations: GAP, GTPase-activating protein; RGS, regulator of G-protein signalling.

protein signaling proteins 21 (*RGS21*) in agreement with the Human Genome Organization (HUGO) Nomenclature Committee. Here we report the isolation and expression pattern of the human *RGS21* gene.

MATERIALS AND METHODS

cDNA library construction and DNA sequencing. The *RGS21* cDNA was cloned from a human fetal library during large-scale cDNA sequencing as described (Li *et al*., 2005). The cDNA library was constructed with a modified pBluescript II SK (+) vector by using human fetal brain mRNA (Clontech). A 0.5 kb DNA fragment containing *Sfi*IA (5′-GGCCATTATGGCC-3′) and *Sfi*IB (5′-GGCCGCCTCGGCC-3′) recognition sites was cloned into *Eco*RI and *Not*I sites of pBluescript SK (+) (Stratagene); the modified vector was then digested with *Sfi*I and the large fragment was excised and purified for library construction. cDNA library was constructed by following the SMART polymerase chain reaction (PCR) cDNA library construction kit protocol (Clontech). After *Sfi*I digestion and cDNA size fractionation, cDNAs >500 bp were ligated into the *Sfi*IA and *Sfi*IB sites of the modified vector and then transformed into DH5α *Escherichia coli* strain. The 96-well R.E.A.L. plasmid kit (Qiagen, Chatsworth, CA, USA) was used to prepare double-stranded plasmid. Sequencing reactions were performed using Big-Dye Primer Cycle Sequencing and Big-Dye Terminator Cycle Sequencing kits (Perkin-Elmer). The complete

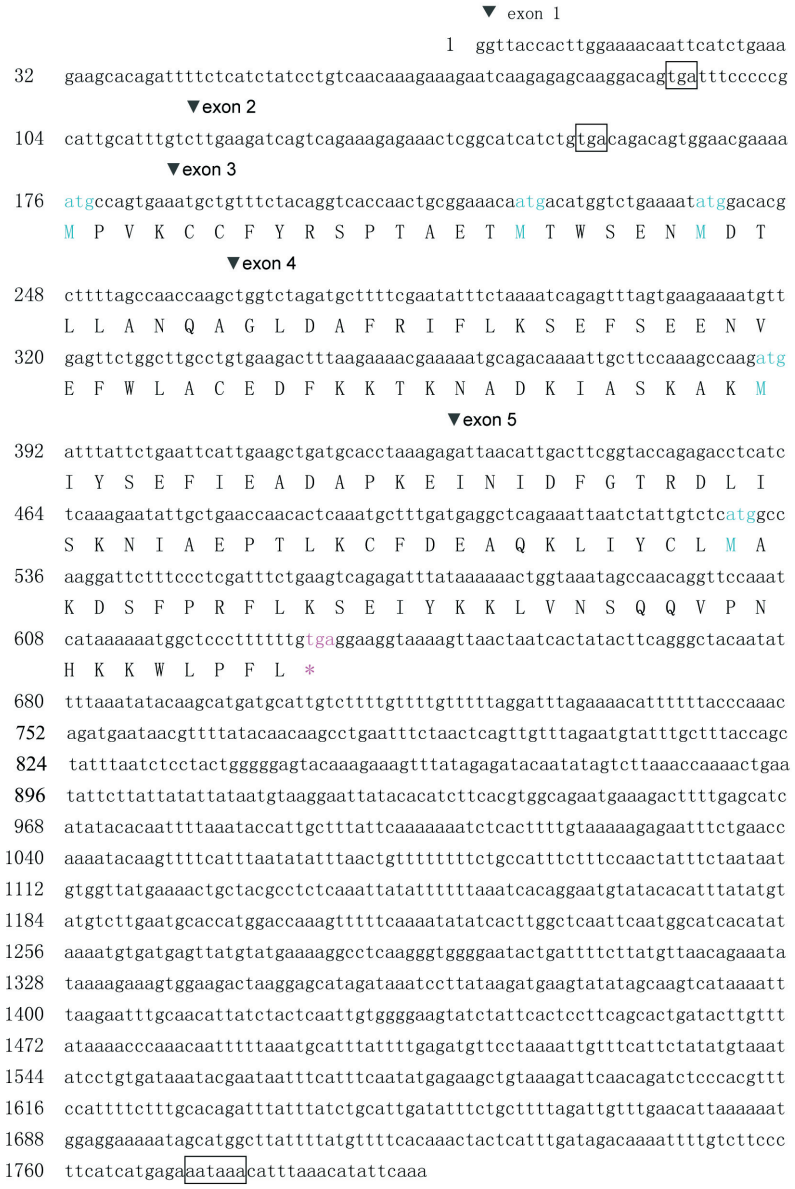

Figure 1. Nucleotide and deduced amino-acid sequence of *RGS21* cDNA.
The open reading frame is from 176 to 631 bp. The in-frame stop codons and possible polyadenylation signal are boxed and the asterisk is positioned to show the stop codon.

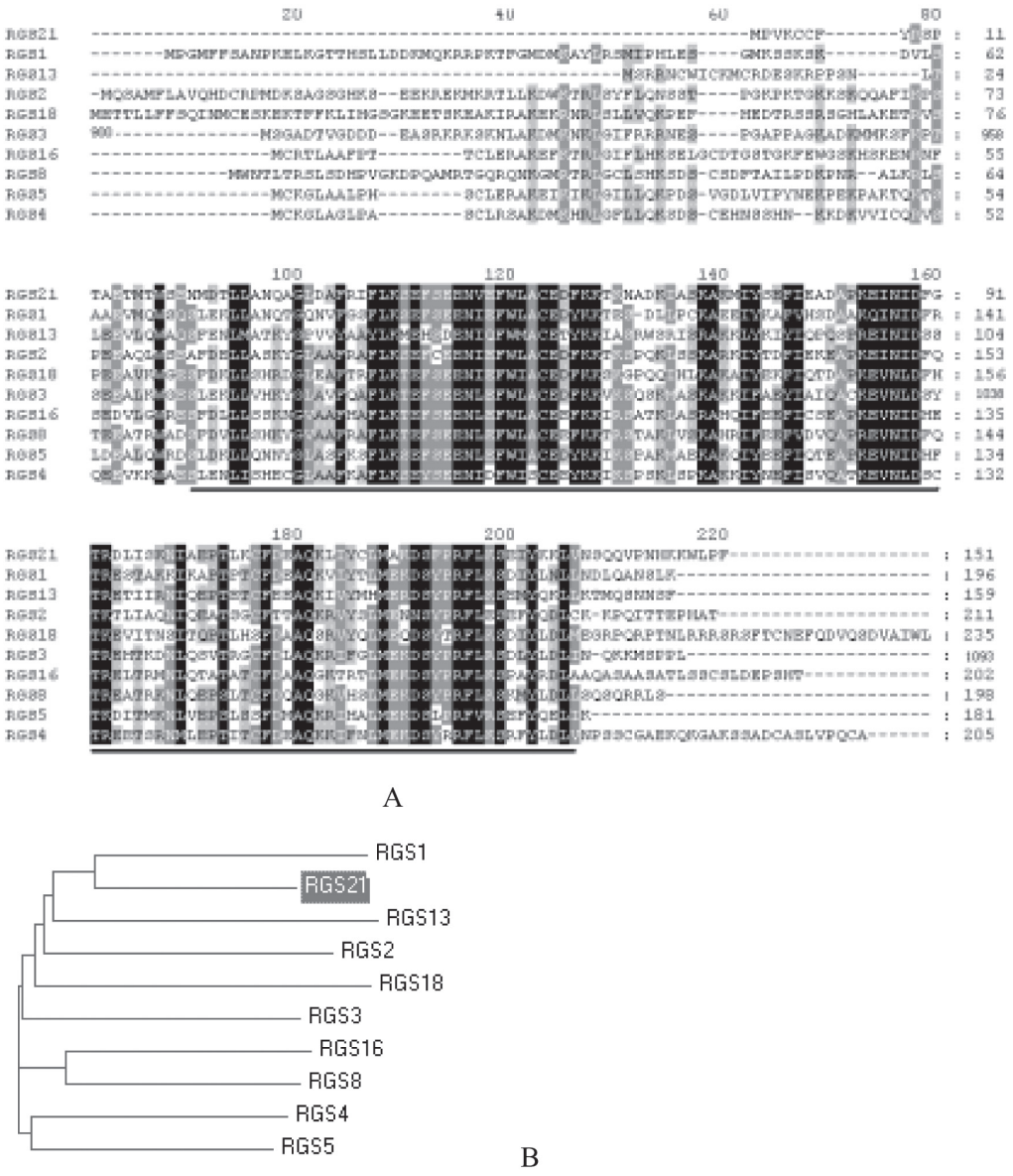

Figure 2. Amino-acid sequence alignment (A) and phylogenetic tree (B) of human B/R4 subfamily members by GeneDoc program.
The location of RGS domain is underlined. Accession numbers of the full-length proteins (excluding C-terminal 900 *aa* of *RGS3*): *RGS1*, NM_002922; *RGS2*, NM_002923; *RGS3*, NM_144488; *RGS4*, NM_005613; *RGS5*, NM_003617; *RGS8*, NM_033345; *RGS13*, NM_002927; *RGS16*, NM_002928; *RGS18*, NM_130782. Clearly, *RGS21* has highly homology with and belongs to the B/R4 subfamily.

sequence was determined and confirmed by primer walking strategy. Sequence assembly was performed with the program Acembly (Sanger's Center).

Sequence analysis. DNA sequence homology searches and comparisons were performed using BLAST-N and BLAST-X on the National Center for Biotechnology Information (NCBI) network service (http://www.ncbi.nlm.nih.gov/blast). The predicted amino-acid sequence of the cDNA was compared against the profile entries to find the occurrence of known profiles (http://www.expasy.ch/pfscan). To identify the chromosomal localization and the gene

Table 1. Nucleotide sequence of exon-intron junctions of *RGS21* cDNA

3′ splice acceptor	exon	size(bp)	5′ splice donor	intron	size (bp)
cDNAend ***ggtta***	1	114	***tttgt***gtgag	1	25872
ttcag***cttga***	**2**	**71**	***gtgaa***gtgag	2	4266
gttag***atgct***	**3**	**77**	***ccaag***gtaag	3	4656
cacag***attaa***	**4**	**1365**	***tcaaa***		

The exon sequence is shown in italic and bold letters. All sequences at the exon-intron junctions are consistent with the AG-GT rule.

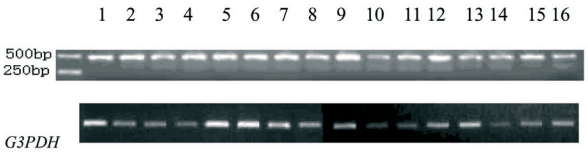

Figure 3. Tissue distribution of human *RGS21* mRNA. PCR analysis of human cDNA for *RGS21* and *G3PDH* (as a control). Prenormalized cDNAs from 16 human adult tissues were purchased from CLONTECH and employed as templates in PCR reactions containing *RGS21*- and *G3PDH*-specific primers as described in the text. Lane: 1, heart; 2, brain; 3, placenta; 4, lung; 5, liver; 6, skeletal muscle; 7, kidney; 8, pancreas; 9, spleen; 10, thymus; 11, prostate; 12, testis; 13, ovary; 14, small intestine; 15, colon; 16, peripheral blood leukocyte. *RGS21* mRNA is expressed ubiquitously in the 16 tissues tested.

structure of human *RGS21*, BLAST-N searching against the human genome was performed. Other analysis was performed by the GeneDoc program.

Assessment of *RGS21* mRNA expression pattern by PCR. Human Multiple Tissue cDNA (MTC) panels (CLONTECH) were used as PCR templates according to the manufacturer's protocol. Thirty-five PCR cycles for human *RGS21* cDNA and 24 for *G3PDH* cDNA (as control) were performed using Taq polymerase (Sangon) at the following program: 0.5 min at 94°C, 1 min at 68°C. The PCR products of human *RGS21* and *G3PDH* cDNA were then electrophoresed on a 2% agarose gel. PCR primers are indicated from 5′ to 3′ as the following: *RGS21* cDNA sense: CGGGATCC ATG CCA GTG AAA TGC TGT TTC TAC AG, *RGS21* cDNA antisense: ATAAGAATGCGGCCGC CAA AAA AGG GAG CCA TTT TTT ATG ATT TG, *G3PDH* cDNA sense: TGA AGG TCG GAG TCA ACG GAT TTG GT, *G3PDH* antisene: CAT GTG GGC CAT GAG GTC CAC CAC. The sense and antisense primers of *RGS21* cDNA span about 460 bp covering the whole open reading frame (ORF).

RESULTS AND DISCUSSION

In this study, a 1.8-kb novel cDNA was isolated and found containing an open reading frame from nucleotide 176 to 631, encoding a 152-amino-acid polypeptide with a predicted molecular mass of 17.7 kDa. It was considered as a full-length cDNA, for two in-frame stop codons (tga) were found at the position 155–157 and 92–94, and one possible polyadenylation signal AATAAA was near the 3′ end of the nucleotide sequence (Fig. 1).

The result of BLAST-N search against the Nr database of GeneBank showed that *RGS21* cDNA was identical, except the 221 bp of the N-terminal, with an RGS-like protein (GeneBank: XM_089307) deduced from the genomic DNA by automated computational analysis. The *RGS21* cDNA was derived from the human DNA sequence on chromosome 1q31.1, spanning about 126 kbp and consisting of four exons. All sequences at the exon–intron junctions are consistent with the AG-GT rule (Table 1). It is located between *RGS18* and *RGS1* gene in a cluster on this chromosome 1q31.1. The cluster contains five RGS genes, including *RGS18, RGS21, RGS1, RGS13* and *RGS2*, all of which belong to the R4 subfamily (Sierra *et al*., 2002). BLASTX analysis showed that the amino-acid sequence of RGS21 shared 58%, 57%, 52%, 54% and 47% identity with human *RGS2, RGS1, RGS8, RGS5*, and *RGS18*, respectively (Fig. 2A). RGS21 has a conserved RGS domain and belongs to the R4/B subfamily by phylogenetic tree analysis (Fig. 2B). In contrast to other R4 subfamily members, it does not have an amphipathic helix like RGS2, thus it represents the smallest known member, which suggests that it serves almost exclusively as a negative regulator of a G protein signaling pathway and does not interact with other signaling pathways. Almost all R4 subfamily members are effective GAPs for both αi/o and αq/11 family members, so may RGS21 (Tesmer *et al*., 1997).

To examine its expression pattern, PCR analysis was performed. The results showed that human *RGS21* mRNA was widely expressed in all 16 tissues we used, which suggests it has general and important physiological roles (Fig. 3). Therefore, further experiments are necessary to clarify the precise roles of RGS21.

Acknowledgements

This work was part of project CHLPP (2004BA711A20) supported by the National Nature Science Foundation of China. The nucleotide sequence reported in this paper has been submitted to GenBank under accession number: AY643711

REFERENCES

Berman DM, Wilkie TM, Gilman AG (1996) *Cell* **86**: 445–452.

De Vries L, Mousli M, Wurmser A, Farquhar MG (1995) *Proc Natl Acad Sci USA* **92**: 11916–11920.

Hollinger S, Hepler JR (2002) *Pharmacol Rev* **54**: 527–559.

Jones PG, Macdonald SG, Cockett MI (2000) *Expert Opin Ther Patents* **9**: 1–14.

Xin Li, Chaoneng Ji, Jiefeng Gu *et al*. (2005) *Mol Boil Rep* **32**: 127–131.

Simon ML, Strathmann MP, Gautam N (1991) *Science* **252**: 802–806.

Tesmer JJG, Berman DM, Gilman AG, Sprang SR (1997) *Cell* **89**: 251–261.

Zheng B, De Vries L, Farquhar GM (1999) *Trends Biochem Sci* **24**: 411–414.

Zhong H, Neubig RR (2001) *J Pharmacol Exp Ther* **297**: 837–845.